

B. Campbell

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/530,112A

1632

DATE: 12/17/98
TIME: 18:24:30

INPUT SET: S30390.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: KENICHI MATSUBARA
6 KOUSAKU OKUBO
7
8 (ii) TITLE OF INVENTION: GENE SIGNATURE
9
10 (iii) NUMBER OF SEQUENCES: 7848
11
12 (iv) CORRESPONDENCE ADDRESS:
13
14 (A) ADDRESSEE: Lyon & Lyon
15 (B) STREET: 633 West Fifth Street
16 Suite 4700
17 (C) CITY: Los Angeles
18 (D) STATE: California
19 (E) COUNTRY: U.S.A.
20 (F) ZIP: 90071
21
22 (v) COMPUTER READABLE FORM:
23
24 (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
25 storage
26 (B) COMPUTER: IBM Compatible
27 (C) OPERATING SYSTEM: IBM P.C. DOS 5.0
28 (D) SOFTWARE: Wordperfect 5.1
29
30 (vi) CURRENT APPLICATION DATA:
31
32 (A) APPLICATION NUMBER: 08/530,112
33 (B) FILING DATE: July 10, 1995
34
35 (vii) PRIOR APPLICATION DATA:
36
37 (A) APPLICATION NUMBER: PCT/JP94/01916
38 (B) FILING DATE: 11 November 1994
39
40 (A) APPLICATION NUMBER: Japanese H5-355504
41 (B) FILING DATE: 12 November 1993
42
43 (viii) ATTORNEY/AGENT INFORMATION:
44
45 (A) NAME: Warburg, Richard J.
46 (B) REGISTRATION NUMBER: 32,327

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/530,112ADATE: 12/17/98
TIME: 18:24:31

INPUT SET: S30390.raw

47 (C) REFERENCE/DOCKET NUMBER: 215/070.
48
49 (ix) TELECOMMUNICATION INFORMATION:
50
51 (A) TELEPHONE: (213) 489-1600
52 (B) TELEFAX: (213) 955-0440
53 (C) TELEX: 67-3510
54
55 (2) INFORMATION FOR SEQ ID NO: 1:
56 (i) SEQUENCE CHARACTERISTICS:
57 (A) LENGTH: 704
58 (B) TYPE: nucleic acid
59 (C) STRANDEDNESS: single
60 (D) TOPOLOGY: linear
61 (vii) IMMEDIATE SOURCE:
62 (B) CLONE: HUMGS00001
63 (ix) FEATURE:
64 (D) OTHER INFORMATION: "N"=unidentified base
65 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
66 GATCTTCAAA CAAGCATCAG CGTTTTCCAG GGCTTCCCAG AGGTCTGTGC GACTAGCCCCG 60
67 TGTCTATCAA AAGTTATTAG AGAGGATGAA GCATTAGCTT GAAGCACTAC AGGAGGAATG 120
68 CACCACGGCA GCTCTCCGCC AATTTCTCTC AGATTTCCAC AGAGACTGTT TGAATGTTTT 180
69 CAAAACCAAG TATCACACTT TAATGTACAT GGGCCGCACC ATAATGAGAT GTGAGCCTTG 240
70 TGCATGTGGG GGAGGAGGGA GAGAGATGTA CTTTTTAAAT CATGTTCCCC CTAAACATGG 300
71 CTGTTAACCC ACTTGCATGC AGAAACTTGG GATGTCACCT GCCTGACATT CACTTTCCAG 360
72 GAGAGGACCC TATCCCCAAA TGTGGAATTG ACTTGCCCTAT GGCCAAGGTC CCTTGGNAAA 420
73 GGGAGCTTCA GTATTTGTGG GGCNTCATA AAACCATGGN TTCAAGNCAA TCCAGCCTCA 480
74 TNGGGNNGGT CCTGGGNACA GTTTTTTGGT AAAGGCCCTT GGCCCAAGNTG GGGGGAATGG 540
75 GCCTCCTTTT TAAGNTTTGG GNTGGAATNG TCTNGCAAAT TGGGGCTCCC ATTTTCNCGGG 600
76 GGTTTGGGGG TTTTNTNGG CTTNCCNGG NNGGAAGGN TGGGTTTGGG GGNTNGGTTN 660
77 CCNTTGGGNG GGCCTGGGNG TTTGATTNA CCCGGNCTT NGGN 704
78
79 (2) INFORMATION FOR SEQ ID NO: 2:
80 (i) SEQUENCE CHARACTERISTICS:
81 (A) LENGTH: 659
82 (B) TYPE: nucleic acid
83 (C) STRANDEDNESS: single
84 (D) TOPOLOGY: linear
85 (vii) IMMEDIATE SOURCE:
86 (B) CLONE: HUMGS00002
87 (ix) FEATURE:
88 (D) OTHER INFORMATION: "N"=unidentified base
89 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
90 GATCTTTAAA ATACACACTC AAATCAAGAA ACTTAAGGTT ACCTTTNTTC CCAAATTTCA 60
91 TACCTATCAT CTTAAGTAGG GACTTCTGTC TTCACAACAN ATTATNACCT TACAGAAGTT 120
92 TGAATTATCC GGTCTGGGTTT TATTGTTTAA AATCATTTC TGCATCAGCTG CTGAAACAAC 180
93 AAATAGGAAT TGTTTTTTATG GAGGCTTTGC ATAGATTCCC TGAGCAGGAT TTTAATCTTT 240
94 TNCTAACTGG ACTGGTTCAA ATGTTGTNCT CTCTTTTAAA GGGATGGCAA GATGTGGGCA 300
95 GTGATGTCAC TTAGGGCAGG GACAGGATAA GAGGGNTTAG GGAGAGAAGA TAGCAGGGCA 360
96 TGGCTGGGAA CCCAAGTCCA AGCATACCAA CACGGAGCAG GCTACTGTCA AGCTCCCCTC 420
97 GGAGGCGGNG CTGGTTCACA GCCAGCTGGC ACCAGNTTT NTNGNGGAAG NCTTTTTTCAA 480
98 ACAGTCTCAG GNAATCCAAT NTGCAAAGAC TTGCTTTNAG NAAAACCCAG NAGTTGAAAG 540
99 GCTCCCAAGN ATTTTAAGGG NACTTNCCAA AACGGGGCCC CNGGNNCCTT TTGGGTTTNG 600

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/530,112ADATE: 12/17/98
TIME: 18:24:32

INPUT SET: S30390.raw

100 GGGNTCAAAA CCCCGGAGGG GTTGGGAAG NTTTTAATTG GNTTTAAAN ATNNNTNTN 659
101
102 (2) INFORMATION FOR SEQ ID NO: 3:
103 (i) SEQUENCE CHARACTERISTICS:
104 (A) LENGTH: 625
105 (B) TYPE: nucleic acid
106 (C) STRANDEDNESS: single
107 (D) TOPOLOGY: linear
108 (vii) IMMEDIATE SOURCE:
109 (B) CLONE: HUMGS00003
110 (ix) FEATURE:
111 (D) OTHER INFORMATION: "N"=unidentified base
112 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
113 GATCTAACTG GGTACCTGAG ATATTTNACA GCTGGACCTA GTTTCACAAT CTGTTGTCTC 60
114 CAGCTCTGCA TATGTCTGGC CAGGGGGCTT CTAGGAAGTA GGTTTCATCT ATCAAATGTC 120
115 TCCTCTGACT TCCTTTTGAA ACTTACTGCT CTTCTGTTTT ATTTTGTTTT GTTTGAAGCT 180
116 CAGAGGGAGA TGGGCAATTG ACAGGGATGC AATCCAGGGT GGGATTTCTT GAGGAAGTTA 240
117 CAAATAAGCT TGTTACAACA TCAAGATAGA TGGAATTGGA AGGATGCTAC CAGGAGAGTA 300
118 CTTACATAGT GCTCAGGAGT TTCTCTCTT AAAATGTTTA CTGCTGAAAG ATGAGCAGGA 360
119 CCAGGGCGTT ATAGGCAGAG CCCTAGCCGA GAAACCTGCT GGCCTCTGCC TGTTTTCATT 420
120 TCCCACTTTT GGTTGTTGTG GCATTACTTT CAGAAATTCG ACTTTCCTGC TTGTCATGAC 480
121 TTTTTTGGCA CACTTGCCAT GACGGGTGTT TCTGNGAACC ATGGAAGTTT TGCGGTAGTG 540
122 CCTCCAGGGG CAGGGGGNAA GGAGNGGTG TANCTGCATT TNGTNCAAAT AAATCCNGCC 600
123 TATTGTTAAT NAACCACTCT TTTGN 625
124
125 (2) INFORMATION FOR SEQ ID NO: 4:
126 (i) SEQUENCE CHARACTERISTICS:
127 (A) LENGTH: 256
128 (B) TYPE: nucleic acid
129 (C) STRANDEDNESS: single
130 (D) TOPOLOGY: linear
131 (vii) IMMEDIATE SOURCE:
132 (B) CLONE: HUMGS00004
133 (ix) FEATURE:
134 (D) OTHER INFORMATION: "N"=unidentified base
135 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
136 GATCTGCCAG GCTGGGGTGT TTTCGGTATC TGCTGTTTAC AGGTCTCCAC TGTAATCCGA 60
137 ATACTTTGCC AGTGCACTAA TCTCTTTGGA GATAAAATTC ATTAGTGTGT TACTAAATGT 120
138 NAATTTTNTT TTGCGGAAAA TACAGTACCG TGTCTGANTT AATTATTAAT ATTNAAAATA 180
139 CTTTCATTCCT TAACCTCTCC TCATTTGCTT TGCCACAGC CTATTCAGTT CCTTTGTTTG 240
140 GCAGGNTTCT GCAAAA 256
141
142 (2) INFORMATION FOR SEQ ID NO: 5:
143 (i) SEQUENCE CHARACTERISTICS:
144 (A) LENGTH: 616
145 (B) TYPE: nucleic acid
146 (C) STRANDEDNESS: single
147 (D) TOPOLOGY: linear
148 (vii) IMMEDIATE SOURCE:
149 (B) CLONE: HUMGS00005
150 (ix) FEATURE:
151 (D) OTHER INFORMATION: "N"=unidentified base
152 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/530,112ADATE: 12/17/98
TIME: 18:24:33

INPUT SET: S30390.raw

153 GATCAAGCTC AAGAATAAGC TGAAATATGG CCAGACTATC AGGCCCATTT NTCTCCCCTG 60
154 CACCGAGGGA ACAACTCGAG CTTTGAGGCT TCCTCCAACT ACCACTTGCC AGCAACAAAA 120
155 GGAAGAGCTG CTCCCTGCAC AGGATATCAA AGCTCTGTTT GTGTCTGAGG AGGAGAAAAA 180
156 GCTGACTCGG AAGGAGGTCT ACATCAAGAA TGGGGATAAG AAAAGGCAGCT GTGAGAGAGA 240
157 TGCTCAATAT GCCCCAGGCT ATGACAAAGT CAAGGACATC TCAGAGGTGG TCACCCCTCG 300
158 GTTCCTTTGT ACTGGAGGAG TGAGTCCCTA TGCTGACCCC AATACTTGCA GAGGTGATTC 360
159 TGGCGGCCCC TTGATAGTTC ACAAGAGAAG TCGTTTCATT CAAAGTTGGT GTAATCAGCT 420
160 GGGGGAGTAA GTGGGATGTN TGCAAAAACC AGAAGGCGGC AAAAGCAGGT ACCTGNNTCA 480
161 NGCCCCGAAC TTTACATCA NCTNTTTCAA GGTNTTNCTT GGNTGAGGNG GAACNTCCAN 540
162 GTTGGGGGTT TTGGGTTTTT TTAAGGGGTT CNTGTTNGCA AGGGGGTTGG GNTTNNTTTA 600
163 NNCTGTNN GNACCN 616
164

165 (2) INFORMATION FOR SEQ ID NO: 6:

166 (i) SEQUENCE CHARACTERISTICS:

167 (A) LENGTH: 615
168 (B) TYPE: nucleic acid
169 (C) STRANDEDNESS: single
170 (D) TOPOLOGY: linear

171 (vii) IMMEDIATE SOURCE:

172 (B) CLONE: HUMGS00006

173 (ix) FEATURE:

174 (D) OTHER INFORMATION: "N"=unidentified base

175 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

176 GATCTCTGCT ATTTTANCCC CCCAAATAAG TTATTTGTCC TTAAAGGTTG GTTACNNATA 60
177 ATACCCCTCA GTAAGATTCC AGTATTAATT TCTGGGCAGT TTGTTCTCTG TATACAATTG 120
178 CAAATGATAA GCATTTTGT GAGTGACCAC CTTTGCAATA TGTTTGTTAA TTNTTCATGT 180
179 TGGGTTCTTT CTGAAATGTA CATCTTTACA TAAAAACCTC ACATTCTACT TGATTTACAC 240
180 TTCCTAGTCT ACATTACATG TGGTTGAAGG TTTTATACAT TCTATATGCT TTTACTAAAT 300
181 ATACAAGATT TACTACTAGA AATTTGGAGA AAGAACACTA ACACATGTAC TTGTGATTTG 360
182 TTCATGTTAT ATTAAACCT TGAGATTTGT GTATTTATGT AGGGGGGGGT ATTGNCCAGG 420
183 NCTGNTGGTT TTTTGCTCCN TGGGGCTATT NTAATAAANC NGGGGTATNG GGTGNTGGC 480
184 CTNGGTTTGG GCCTAAANTT GGATATATNT GGGGTTCCCT NGNTTTTACC AAAATNGNTT 540
185 TTGGTTGGTA GGGTTNTTAT TGGACCCNNT CCTGGGCCTG GGATAATNTG GCNGNTTCCC 600
186 NGGANAAAAAN NCCCN 615
187

188 (2) INFORMATION FOR SEQ ID NO: 7:

189 (i) SEQUENCE CHARACTERISTICS:

190 (A) LENGTH: 608
191 (B) TYPE: nucleic acid
192 (C) STRANDEDNESS: single
193 (D) TOPOLOGY: linear

194 (vii) IMMEDIATE SOURCE:

195 (B) CLONE: HUMGS00007

196 (ix) FEATURE:

197 (D) OTHER INFORMATION: "N"=unidentified base

198 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

199 GATCAAAACCA AAGGAAAAGT GTTGCTAGAG AAAATTGGGG AAAAGGTGAA AAAGAAAAAA 60
200 TGGTAGTAAT TGAGCAGAAA AAAATTAATT TATATATGTA TTGATTGGCA ACCAGATTTA 120
201 TCTAAGTAGA ACTGAATTGG CTAGGAAAAA AGAAAACTG CATGTTAATC ATTTTCCCTAA 180
202 GCTGTCCTTT TGAGGCTTAG TCAGTTTATT GGGAAAAATG TTAGGATTAT TCCTTGCTAT 240
203 TAGTACTCAT TTTATGTATG TTACCTTCA GTAAGTTCTC CCCATTTTAG TTTTCTAGGC 300
204 TGAAAGGATT CTTTTCTACA TTATACATGT GTGTGTTCAT ATTTGGCTTT TGCTATATAC 360
205 TTTACCTTCA TTGTTAAATT TTTGTATTGT ATAGTTNCTT TGGNGGTATC TTTAAACCCCT 420

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/530,112ADATE: 12/17/98
TIME: 18:24:33

INPUT SET: S30390.raw

206 ATTTTGTGAAA ACCAACCTTG GCTTTGNTAA NCATTTGGGC CGCTTGGGTA NGTCCGGACC 480
207 TNNCCTTTNC CCCNAGGGCC TGTCAGGAGN GCCGGNTTN CCGGGGNGTT GGNCCCGNNG 540
208 CCTTCCCGAT TTTTNGGGGT TGGGTCCTTN GGTNCCCGC NTCTGTANGG GGTNCNCCTT 600
209 TTNNNCCN 608

210

211 (2) INFORMATION FOR SEQ ID NO: 8:

212 (i) SEQUENCE CHARACTERISTICS:

213 (A) LENGTH: 606

214 (B) TYPE: nucleic acid

215 (C) STRANDEDNESS: single

216 (D) TOPOLOGY: linear

217 (vii) IMMEDIATE SOURCE:

218 (B) CLONE: HUMGS00008

219 (ix) FEATURE:

220 (D) OTHER INFORMATION: "N"=unidentified base

221 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

222 GATCGCTTCC TAGAAATAAG CAACACCTCT CCCAAAAAGC AGCCCACAAG GCAGGGGCCA 60
223 GCAGCCCAGC CATCACTCAT CTTTNAGGAA ATNAGTTGGT AGCCTCTGTG CACTGTTTGG 120
224 TGGCCACATC ANNGGTGATG TCCTGTTCAC ATACCTGCTT GTATTTAAAG CCCTCAGTCT 180
225 GTCCTGTTGT GTGGGGCGAA GTGATGGACT CTGCCAGGTG GACATGCTGT GGGTGGATGT 240
226 TCCCGGCGTG TGCCGGCCTG AATGGACAGG GGTCACTTCA CAAGCATGTC AGGGAAAATC 300
227 ACTGTCACAC AATTCCAATG GATTTTGTGC TCTTTTGTGA AAAAAAAAAAT TCTTTAGCGT 360
228 AAACCATGNA TTTTTTTTC AATGTAGNCCC TTGGGGANTG AANTGAAATT TTGGGCTTCT 420
229 TCANATNCGN AAAATNAAAT TTTTACCCCT GAGGGGGGGA GCCCCTTCTT GAAAGAAGGT 480
230 NTGGGGCCAA AGCCCTTTTA ATGNTTGCTG GCCNTTGNTG GTTTTNANNG TCCANTTTGC 540
231 TTGGGGCGAA NGCCGNNTG ANAAGGTGG GTTTCNCTGG NGGNTTTAAG GNGGTGGTTT 600
232 GNTTTN 606

233

234 (2) INFORMATION FOR SEQ ID NO: 9:

235 (i) SEQUENCE CHARACTERISTICS:

236 (A) LENGTH: 606

237 (B) TYPE: nucleic acid

238 (C) STRANDEDNESS: single

239 (D) TOPOLOGY: linear

240 (vii) IMMEDIATE SOURCE:

241 (B) CLONE: HUMGS00009

242 (ix) FEATURE:

243 (D) OTHER INFORMATION: "N"=unidentified base

244 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

245 GATCTTCANN NGTTAAGCAC TTGCTCTNAA GATTAATAAT CCTTTTCTTT TTAAGGTAA 60
246 GGGTGTGTAC GTATGGCAGT GATGCTATG TTGAGATTAA CTTATGTATT GAGGAAAATT 120
247 TGAAGTTTAT TTTTTCGATG AATAAGGCTG TCAAATNATT TAGTATAGAT TAATGACATC 180
248 TTTTTTAGAA ATATTAAAGT GAGTATTCCT CATTATGTCA TCATTTCTGA TAATTAGAGT 240
249 GCTAATTTGA ATGTTAGATA ATGNTTCCAC ATCTATACCT ATTTCTTTCT AGGGCACTTC 300
250 TGACCTGGG GCTTGGGGAT GGCCTTTAGG CACAAGTAGT GTCTGTGTTA AGTTCACATA 360
251 ATGTGTATTT AATGAGAAAC ATTCCNATGT AAAAATGTGT GTATGTGAAC GTATGCNNAC 420
252 ATTGTTATTG TGCACCGTA CATTGTGAAG AAGTAGTTTN GAAATTTTGT AANGCACAAC 480
253 CCTTAANGNG GTGTGGAGTT ATTAAANTGN TGTAGGCNCA AATGTAATGT TTAGCCTATA 540
254 AAAGGCCCTC CTATTGTCCN TNGGCAAGGC TTTGNCTT GNAANTAAAN CCCGTNTTTG 600
255 TTTAAA 606

256

257 (2) INFORMATION FOR SEQ ID NO: 10:

258 (i) SEQUENCE CHARACTERISTICS: